Probabilistic Risk Analysis for Use in Human Health Risk Assessment

Forecast: Cancer Risk

Frequency Chart

10,000 Trials

Mean = 1.17E-6
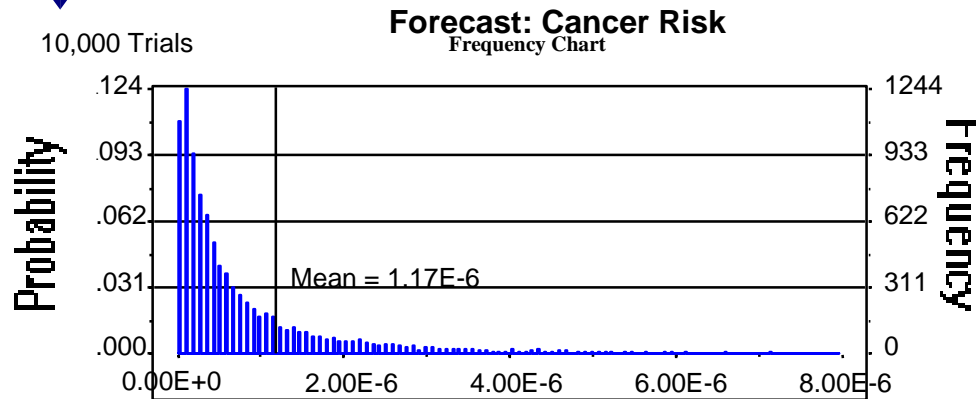
$$\text{Risk} = \frac{Cs \times IR \times 1E\text{-}06 \times EF \times ED \times SF}{BW \times AT}$$

Captain Brian L. Sassaman
San Antonio Air Logistics Center
Environmental Management
October 20, 1999
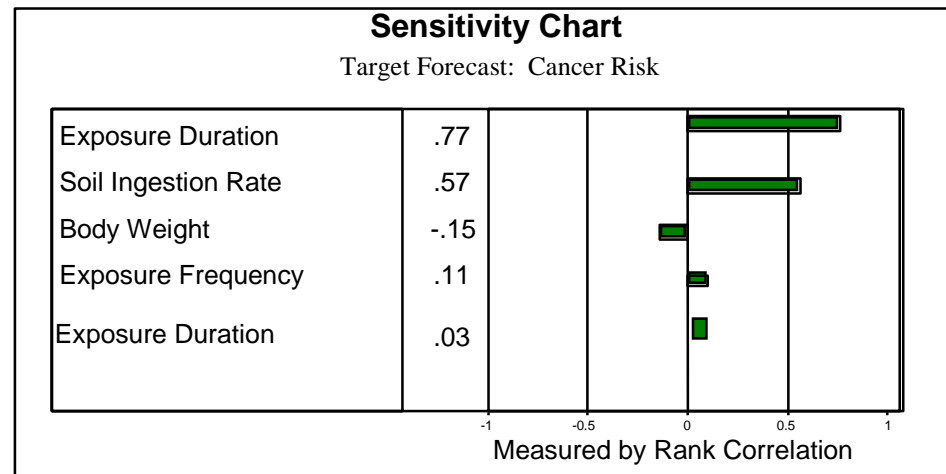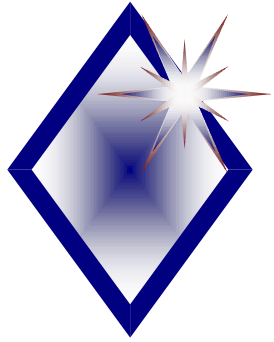
Sensitivity Chart

Target Forecast: Cancer Risk

|		
Exposure Duration	.77	
Soil Ingestion Rate	.57	
Body Weight	-.15	
Exposure Frequency	.11	
Exposure Duration	.03	

Measured by Rank Correlation

Objectives for this Session

- Overview

- Deterministic vs. Probabilistic Modeling

- Principles for use in MC Simulation

- On-going PRA activities

- Conclusion
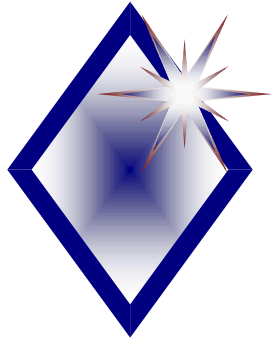
Issues at Hand

- USAF need for state-of-the-art approach to determine potential health risks
  - Protective of human health and the environment
  - Cost-effective and timely
- Approaches
  - Risk Screening
  - Deterministic Risk Assessment
  - Probabilistic Risk Assessment
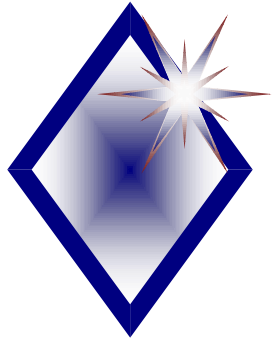
Issues at Hand (cont.)

- EPA's current risk assessment methodology expresses health risks as a single numerical or "single-point" estimates of risk.

- Provides little information about uncertainty and variability surrounding risk estimates.

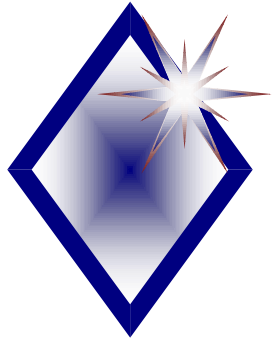- Recent EPA guidance recommends developing "multiple descriptors" of risk.
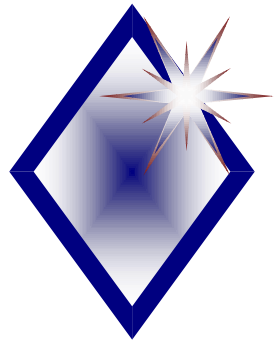
What is Risk?

◆ Risk--the probability of an adverse outcome

An environmental risk
is the probability of an
adverse health effect as
a result of exposure to
a hazardous substance.

Why Model Risks?

- Generally, modeling is performed to:
  - better understand a system
  - make predictions
- Specifically, risk modeling is necessary because:
  - acceptable risk levels are not measurable: toxicology and epidemiology
  - direct sampling is not feasible

> All models are wrong and some are useful (George Box)
>
> All models are wrong and some are useless (Box corollary)
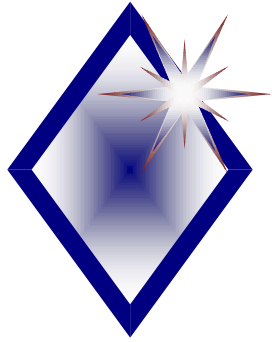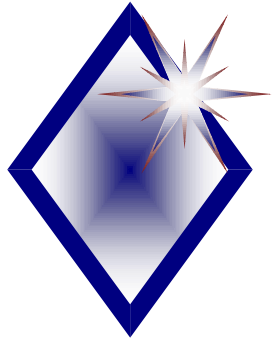
What is Risk Assessment?

An evaluation of the potential health effects of individuals or populations exposed to hazardous materials or situations.

◆ Nature and amount of hazard

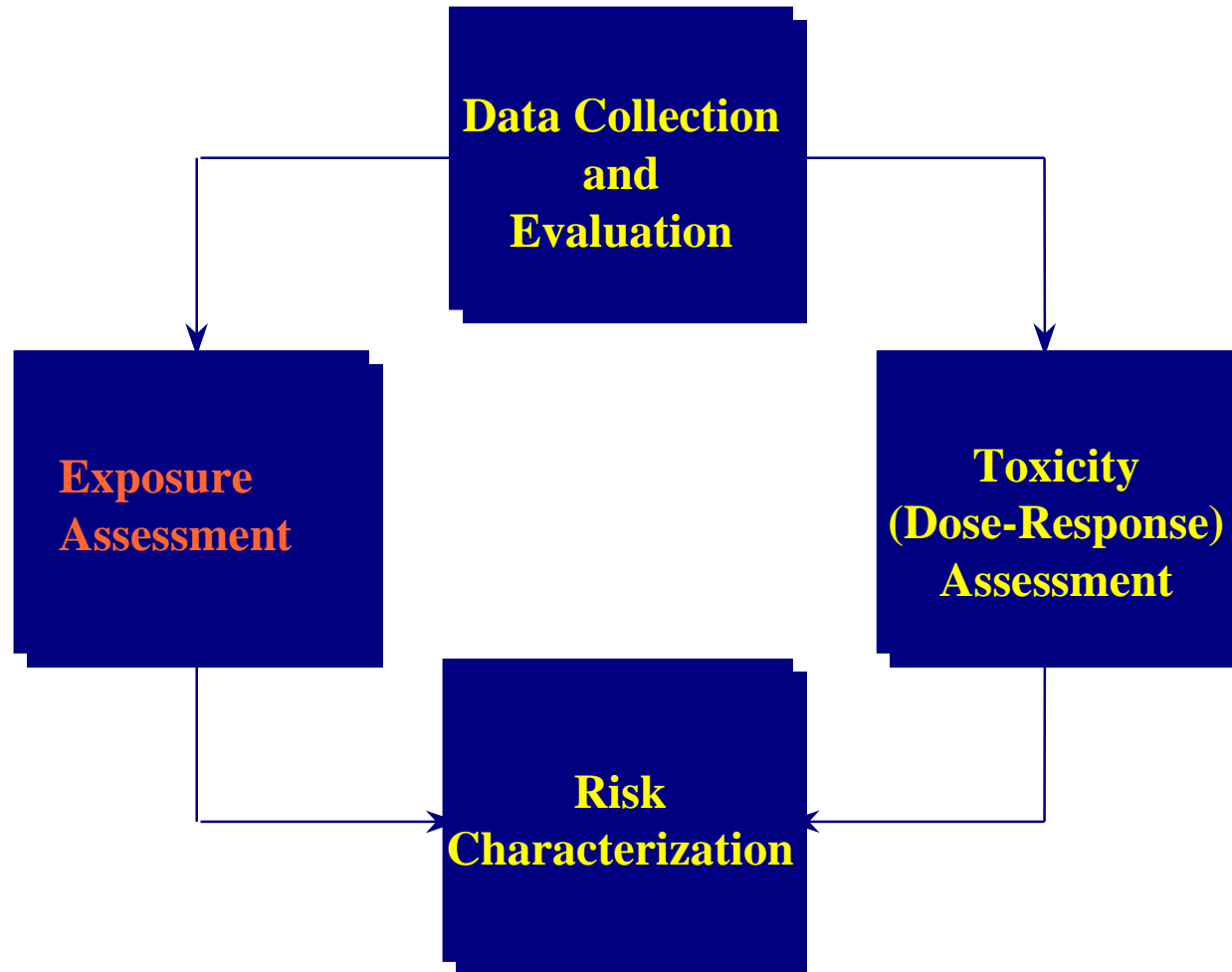◆ Expected health effects

◆ Characteristics of the exposed population
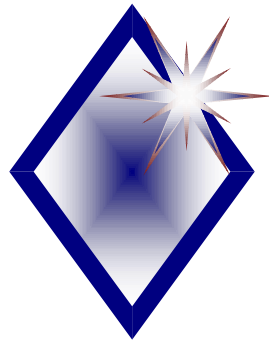
To the extent risk assessment is precise, it is not real

To the extent risk assessment is real, it is not precise

(adapted from Albert Einstein)

Risk Assessment Paradigm

Data Collection and Evaluation

Exposure Assessment

Toxicity (Dose-Response) Assessment

Risk Characterization

Uncertainty: Should we care about uncertainty in the RA?

- Uncertainty about data
  - Data gaps
  - Do the data gaps significantly affect the risk estimate
  - Is new information being developed that might affect the estimate
- Uncertainty about adverse effects
  - Scientific consensus regarding toxicity
  - Will adverse effects occur at this site
- Uncertainty about exposure
  - Will significant exposures occur
 - Probability of significant exposure
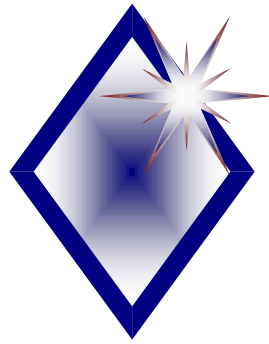
Sources of Uncertainty in Exposure Assessment

- Uncertainty regarding missing or incomplete information to fully define exposure and dose (Scenario Uncertainty)

- Uncertainty regarding some parameter (Parameter Uncertainty)

- Uncertainty regarding gaps in scientific theory required to make predictions on the basis of casual inferences (Model Uncertainty)
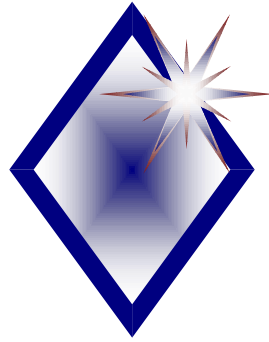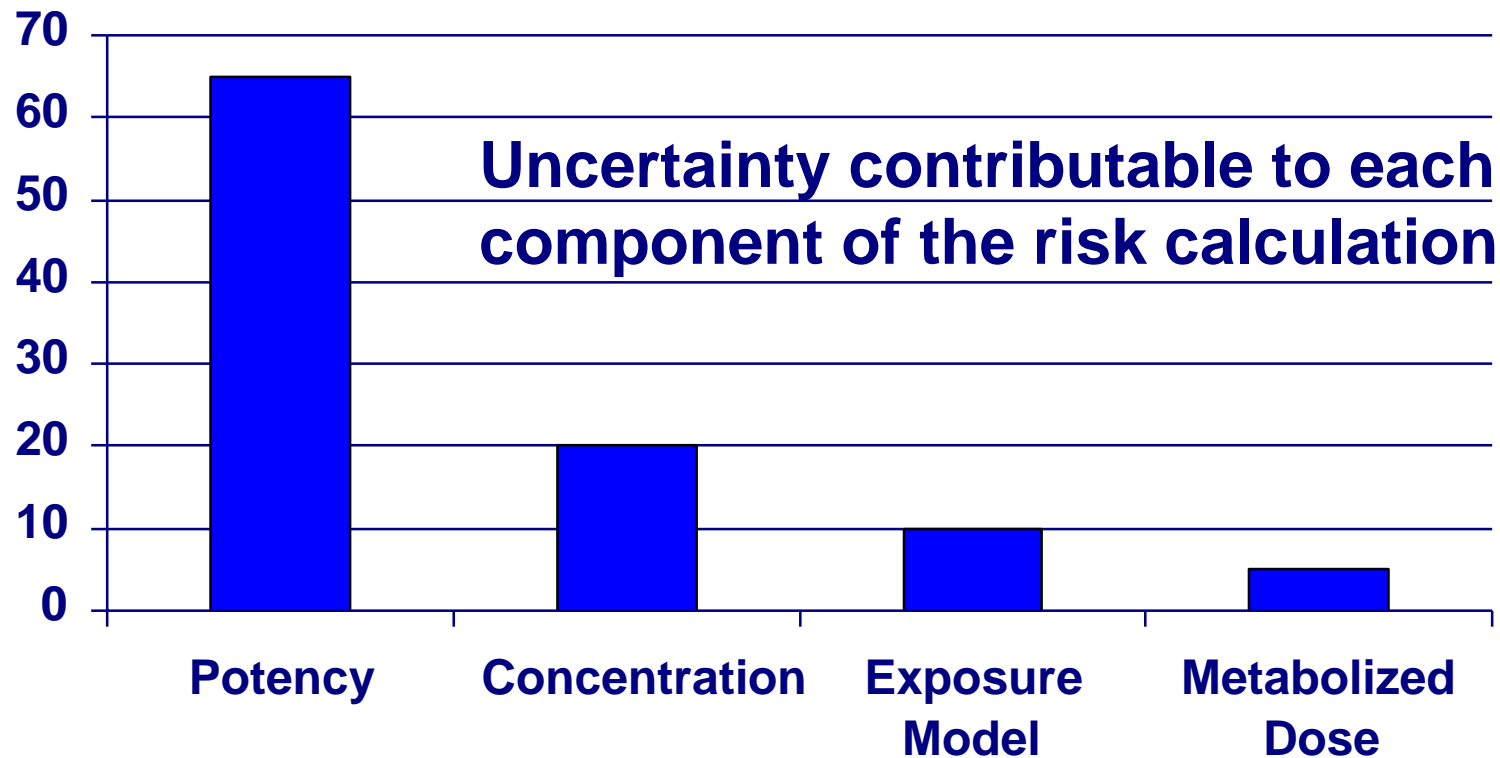
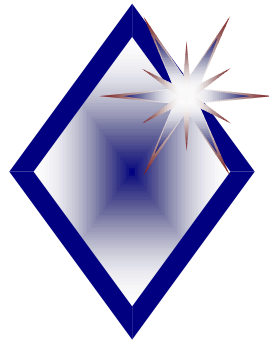# Sources of Uncertainty in Toxicological Measurements

◆ Limited sample size

◆ Measurement endpoints

◆ Extrapolation across species

◆ High-to-low extrapolation

◆ The absence or presence of sensitive populations

Sources of Uncertainty in Toxicological Measurements

Uncertainty contributable to each component of the risk calculation
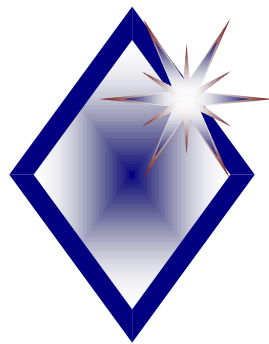
Background - Traditional Risk Calculation

- Risk = Exposure x Toxicity
- For Carcinogens:
  - excess cancer risk = LADD x CSF
    - LADD is lifetime average daily dose
    - CSF is cancer slope factor
- For Non-Carcinogens:
  - hazard quotient = ADD/RfD
    - ADD is average daily dose
    - RfD is reference dose
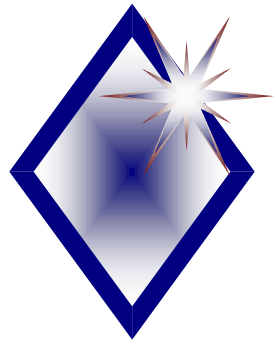
Objectives for this Session

- Overview
- Deterministic vs. Probabilistic Modeling
- Principles for use in MC Simulation
- On-going PRA activities
- Conclusion

Deterministic Approach

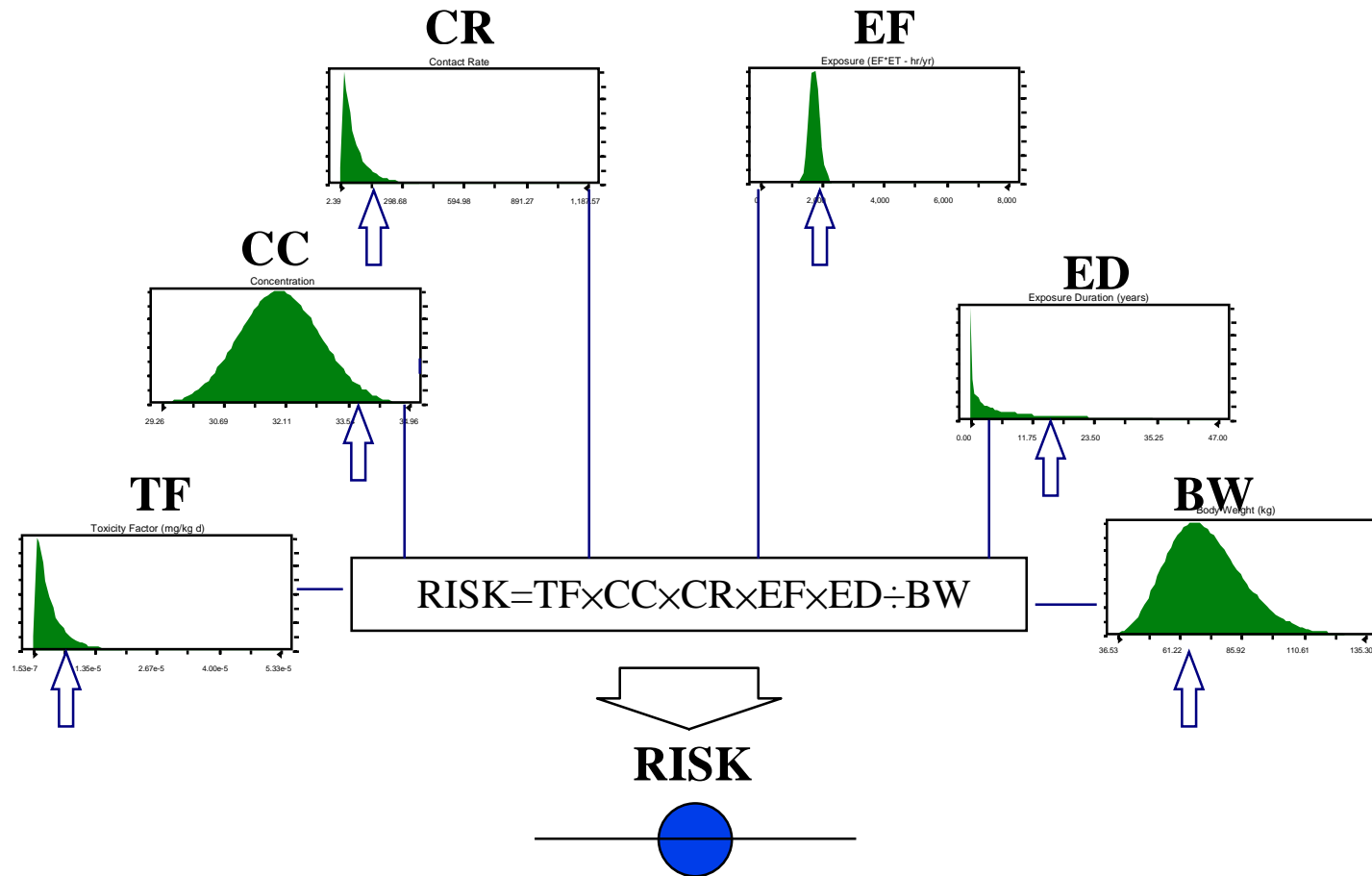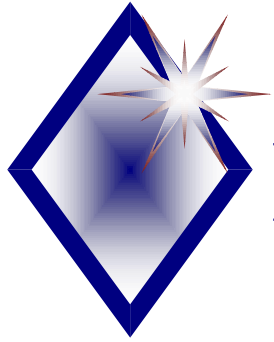**CR**

Contact Rate

2.39   298.68   594.98   891.27   1,187.57

EF

Exposure (EF*ET - hr/yr)

0   2,000   4,000   6,000   8,000

CC

Concentration

29.26   30.69   32.11   33.54   34.96

ED

Exposure Duration (years)

0.00   11.75   23.50   35.25   47.00

TF

Toxicity Factor (mg/kg d)

1.53e-7   1.35e-5   2.67e-5   4.00e-5   5.33e-5

BW

Body Weight (kg)

36.53   61.22   85.92   110.61   135.30

$$RISK = TF \times CC \times CR \times EF \times ED \div BW$$
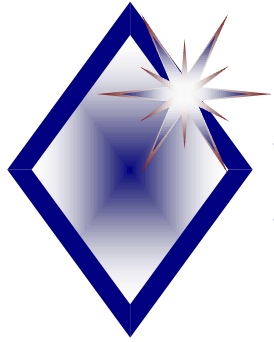
RISK

Example: Drinking Water

$$\text{Risk} = \frac{\text{Cw x IR x EF x ED x CSF}}{\text{BW x AT}}$$

where Cw = concentration in water (ug/L)
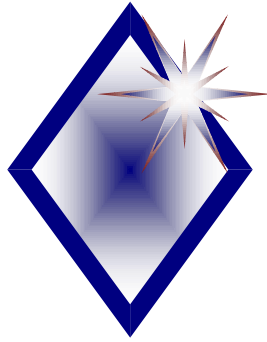
IR = ingestion rate (L/day)

EF = exposure frequency (days/yr)

ED = exposure duration (yrs)

BW = body weight (kg)

AT = averaging time (yrs)

CSF = Cancer slope factor (inverse mg/kd/day)

Exposure from Contaminated Drinking Water, RME

Residential Exposure Scenario

Cw = 100 mg/L   Dose = 1.22 mg/kg/day

IR = 2 L/day   <u>Risk = 3.66E-02</u>

EF = 350 days/yr

ED = 30 years

BW = 70 kg

AT = 70 yrs (24,500 days)

CPF = 3.0E-02 inverse mg/kg/day
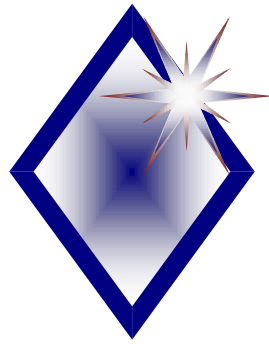
Deterministic Approach to Risk Assessment

- Deterministic Modeling: Approach where a unique value is assigned to each input variable (BW-70 kg) and the output (exposure) is a point estimate (hazard quotient, hazard index, or excess cancer risk)

  - Risk = Toxicity * Exposure = 2 * 5 = 10

- Modeling is "easy" to perform and the prediction is conservative in the protection of human health

- Risk assessors do not need to be versed in the science of probability (reluctant to change).
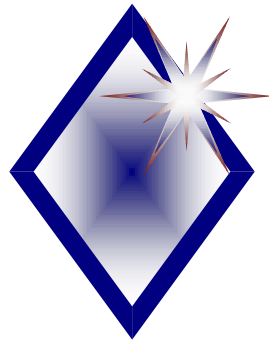
Deterministic Approach to Risk Assessment

◆ Modeling is "easy" to perform; protective of health

◆ RA/RM done need to be versed in probability

◆ Provides little information on range of risks within a given population

◆ Degree of conservatism not easily defined

◆ Provides little information on the uncertainty associated with the risk estimate
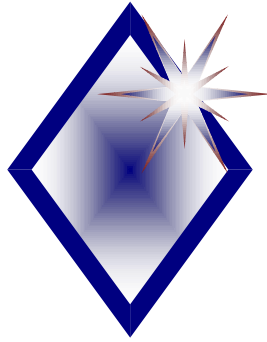
Why Conduct Probabilistic Risk Assessments?

◆ Cost of site cleanups

◆ Demand for credibility

◆ Scientific consensus

◆ EPA's uncertainty analysis guidance

◆ Technical feasibility

◆ Regulatory reform
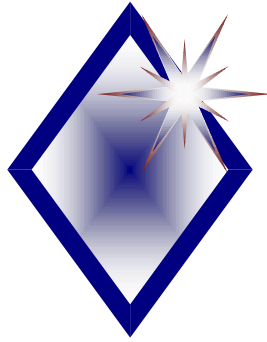
When is a Probabilistic Approach Appropriate?

◆ When screening level calculations indicate a potentially unacceptable concern or risk

◆ To support assessment of the value of collecting additional information

◆ When remediation may result in high costs if uncertainty and variability are not known

◆ When it is necessary to establish the relative importance (rank) of exposures, exposure pathways, sites, or contaminants for further investigation
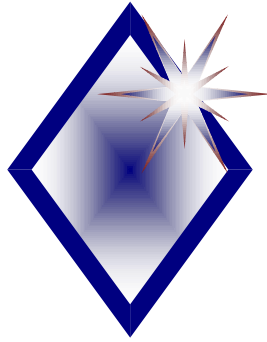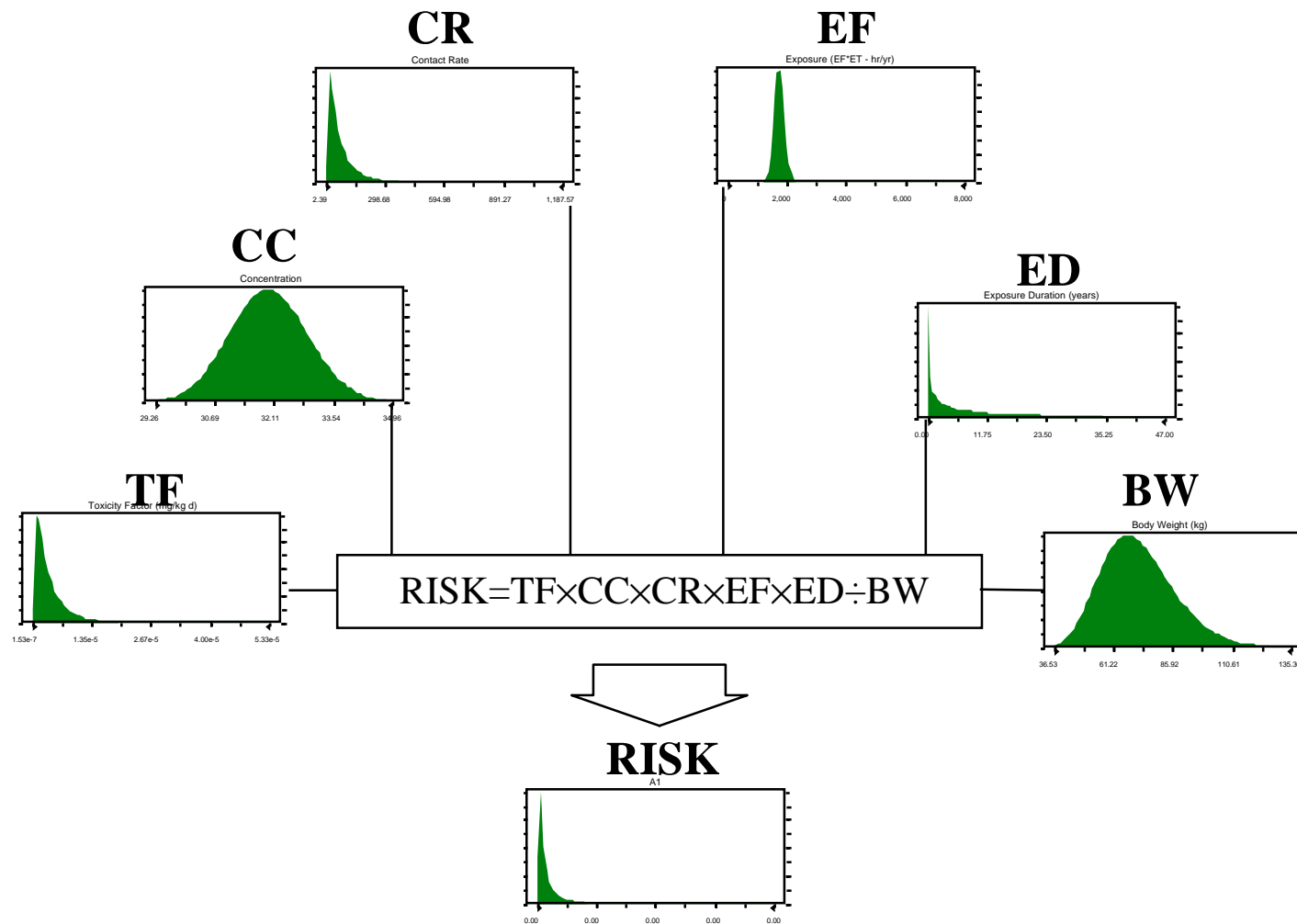
When is a Probabilistic Approach NOT Appropriate?

◆ When screening-level deterministic calculation indicates that exposures are small/acceptable

◆ When there is little uncertainty or variability in the population

◆ When the cost of averting the exposure is smaller than the cost of probabilistic analysis

◆ When there is low trust and high concern about the level of exposure or risk at a site
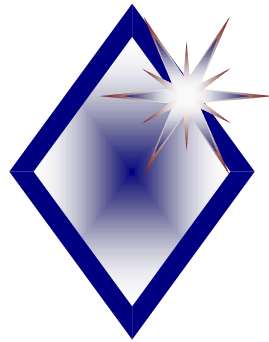
◆ Consequence of a "wrong" prediction is negligible

Probabilistic Modeling Approach

CR

Contact Rate

2.39   298.68   594.98   891.27   1,187.57

EF

Exposure (EF*ET - hr/yr)

2,000   4,000   6,000   8,000

CC

Concentration

29.26   30.69   32.11   33.54   34.96

ED

Exposure Duration (years)

0.00   11.75   23.50   35.25   47.00

TF

Toxicity Factor (mg/kg d)

1.53e-7   1.35e-5   2.67e-5   4.00e-5   5.33e-5

BW

Body Weight (kg)

36.53   61.22   85.92   110.61   135.30

$$RISK = TF \times CC \times CR \times EF \times ED \div BW$$
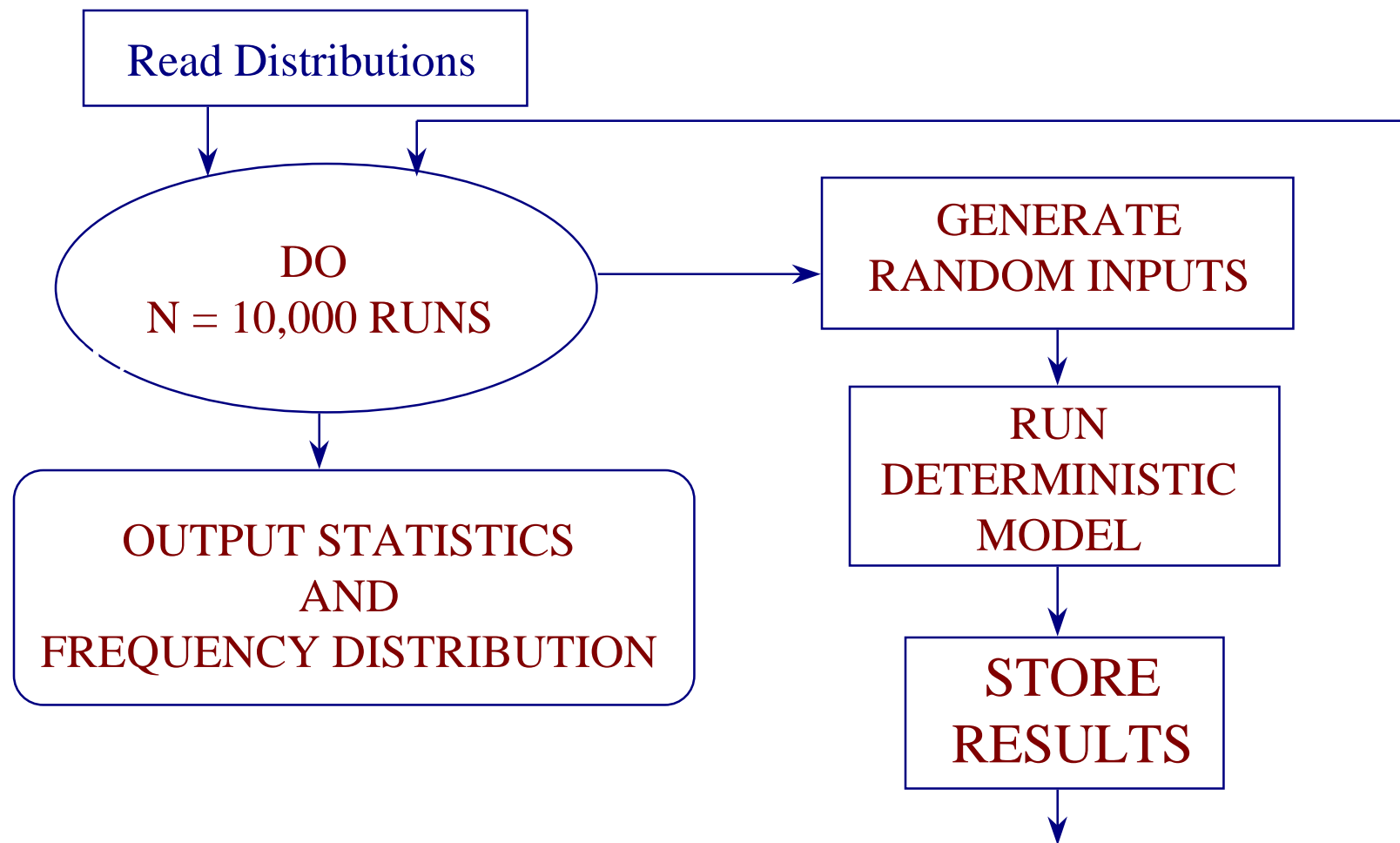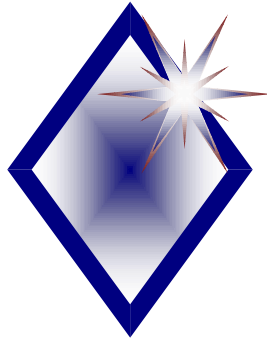
RISK

A1

0.00   0.00   0.00   0.00   0.00
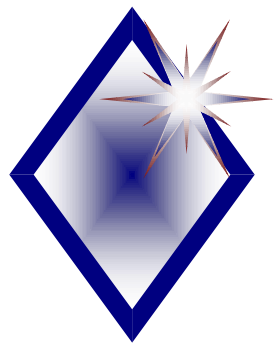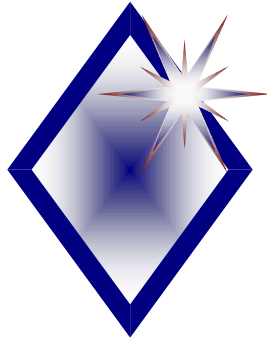
Probabilistic Approach: Definitions

◆ Probabilistic Modeling: Approach where a frequency distribution is assigned to each input variable and the output assumes the form of frequency distribution.

    ◆ Ex: Risk = Toxicity * Exposure = [2,6] * [1,5] = [2,30]

◆ Monte Carlo Simulation: A probabilistic technique by which a prediction is calculated repeatedly using randomly selected "what-if" trials. The results of numerous trials are plotted to represent a frequency distribution of possible outcomes and the likelihood of each such outcome.

Steps in a Probabilistic (Monte Carlo) Approach

Read Distributions

DO
N = 10,000 RUNS

GENERATE
RANDOM INPUTS

RUN
DETERMINISTIC
MODEL

OUTPUT STATISTICS
AND
FREQUENCY DISTRIBUTION

STORE
RESULTS

Probabilistic Approach: Definitions

◆ <u>Variability (Type A Uncertainty)</u>: a characteristic of the population that represents heterogeneity or diversity in a well-characterized population which is usually <u>not</u> reducible through further measurement or study but may be better characterized.

  ◆ For example, different people in a population have different body weights, no matter how carefully we weigh them (physical, chemical, and biological phenomenon).
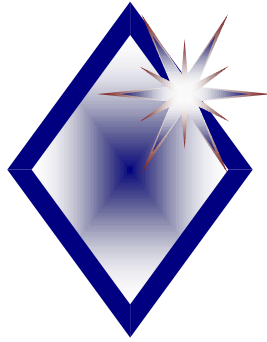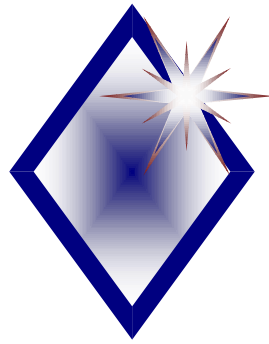
Probabilistic Approach: Definitions

- **Uncertainty (Type B Uncertainty)**: a characteristic of the analyst that represents ignorance about a poorly characterized phenomenon (i.e., specific factors, parameters, or models) which is reducible through further measurement or study.
  - Uncertainty due to availability of data that arises from limitations of study design and analytical technique
  - Uncertainty due to application of the data to nonsampled populations such as between the sampled population and the assessment's target population.
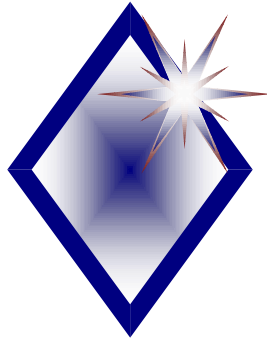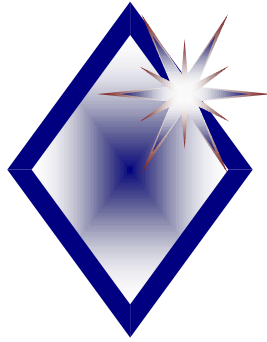
Probabilistic Approach: Input Distributions

◆ The results of a Monte Carlo analysis are only as accurate as the input distribution used

◆ Input distributions should be based on relevant and representative measurement data

◆ Preliminary assessment of distribution based on prior knowledge of data/exploratory data analysis

◆ Commercially available software packages

Probabilistic Approach: Input Distributions

◆ Evaluate the appropriateness of distributions published in the literature, since:

   ◆ Data may be outdated

   ◆ Data may have been developed for a different purpose (e.g., concentrated on hot spots)

   ◆ Data may have been developed for a different exposure population or geographic region

Probabilistic Approach: Aggregation of Distributions from Multiple Sources

◆ Issue in MC simulation is determining how and when to aggregate data from multiple sources into a single input distribution

◆ Multiple sources may include:

  ◆ Data from separate studies that measure common parameters

  ◆ Similar studies conducted by different investigators

 ◆ Different values of input parameters calculated using different methods or models
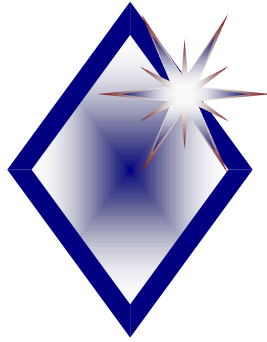
Probabilistic Approach: Aggregation of Distributions from Multiple Sources

◆ To arrive at an accurate determination of risk, make sure that data from various sources <u>appropriately represent the input of interest and the underlying assumptions of risk assessment model</u>
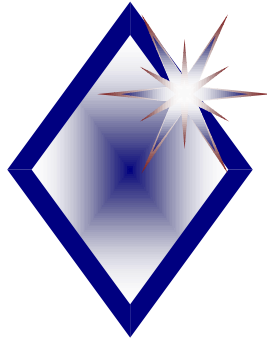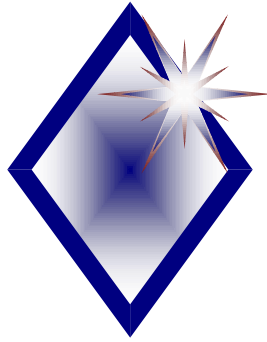
Probabilistic Approach: Aggregation of Distributions from Multiple

- ◆ Example - Octanol-Water Partition Coeff (Kow)
  - ◆ Wide range of data available for PDF for Kow
    - ◆ Data collected using a variety of different measurement methods
  - ◆ Uncertainty in this parameter should not be characterized by including all available data
  - ◆ Most appropriate measurement of technique should be determined, and the uncertainty for Kow should be based on the available data using the selected measurement method
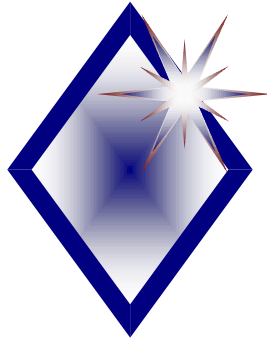
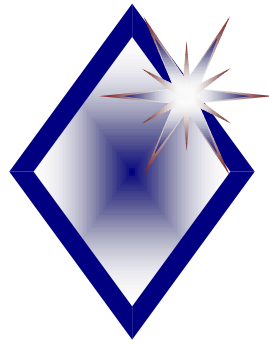# Probabilistic Approach: Information at the Tails of the Input Distributions

◆ May not be as good as the quality of information at the central values of the distributions

◆ Can end up driving the tails of the output distributions (e.g., 90th or 95th percentiles) in the MC analysis

◆ Risk analysts need to pay attention to help ensure the reliability of the estimates at higher percentiles of the output distributions

Probabilistic Approach: Information at the Tails of the Input Distributions

- Example
  - When collecting soil samples for a visually impacted area, a disproportionately large number of samples are often collected from "hot spots"
  - Inclusion of a disproportionate number of "high end" samples when developing an input distribution will bias the distribution parameters
  - If a limited number of samples are collected, the high end data will skew the distribution upwards, and will drive up the tails of the output distribution

Probabilistic Approach: Input Distributions

◆ Distributions are Available On:

|	
Body Weight	Recreational Fishing
Surface Area	Soil Ingestion
Residential Duration	Dermal Adhesion
Occupational Duration	Inhalation Rate*
Tap Water Ingestion*	Dietary Intakes
Morbidity*	

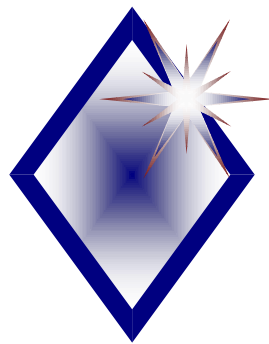*EPA evaluating these parameters for Exposure Factors Handbook
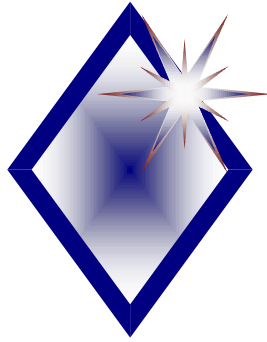
Probabilistic Approach: Correlation

◆ When the values of two variable depend upon one another in part or whole, the variables are considered correlated

◆ Ignoring correlations for sensitive parameters will affect the risk estimate

◆ Example:

  ◆ body weight and and surface area (larger people have more skin surface area)
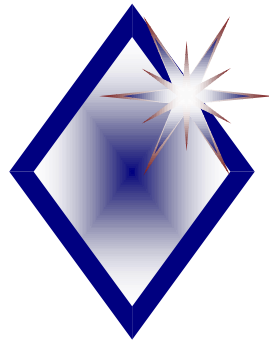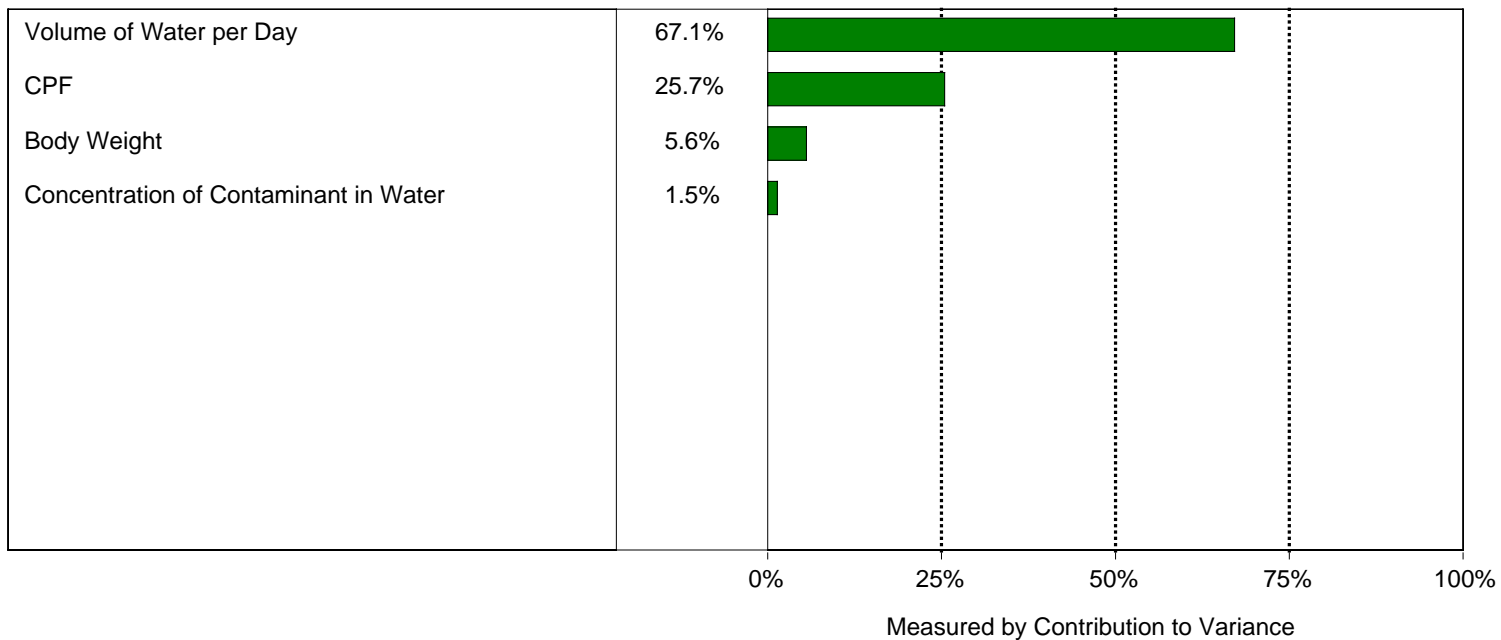
Probabilistic Approach: Sensitivity Analysis

◆ Attempts to provide a ranking of the model's input assumptions with respect to their contribution to model output

◆ Sensitivity analysis attempts to answer the following question:

  ◆ Which exposure factors influence risk most strongly?

  ◆ Which exposure factors can be studied?

  ◆ In which direction should the risk assessment go?
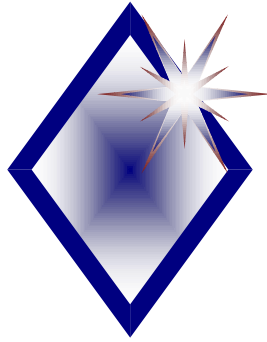
Example of Sensitivity Analysis Output
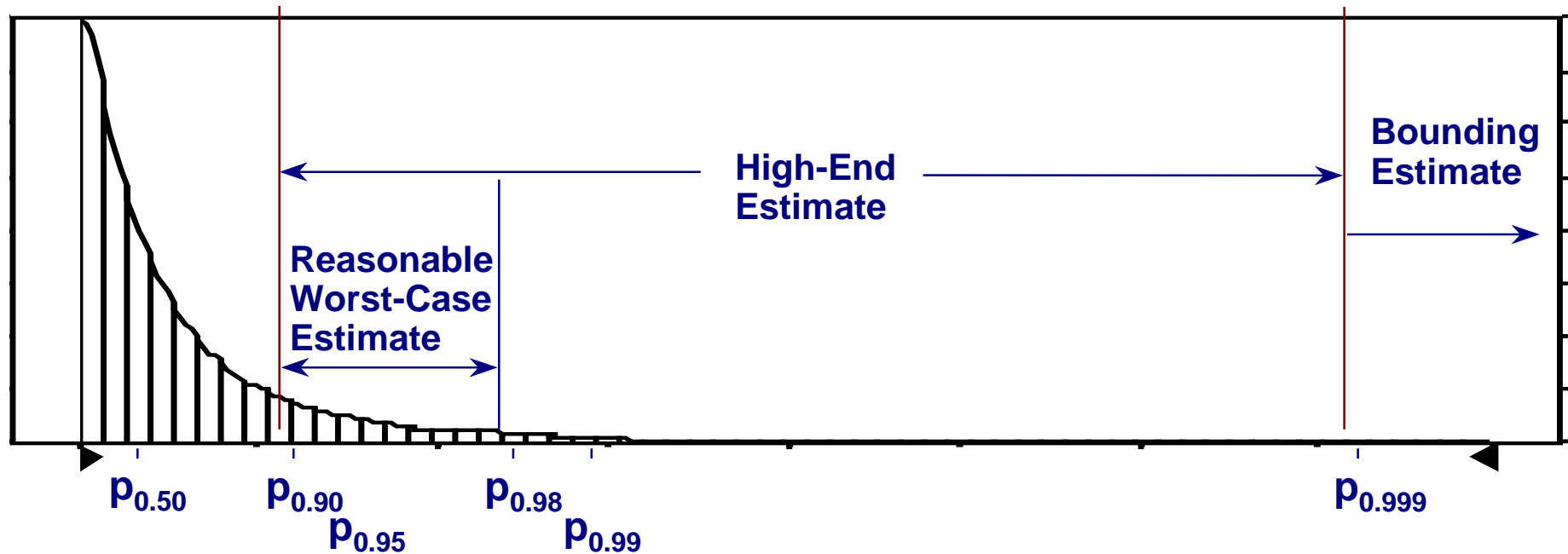
Sensitivity Chart

Target Forecast: Risk Assessment

Volume of Water per Day	67.1%
CPF	25.7%
Body Weight	5.6%
Concentration of Contaminant in Water	1.5%

0%   25%   50%   75%   100%
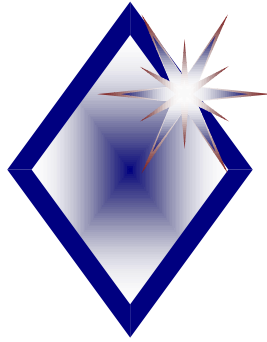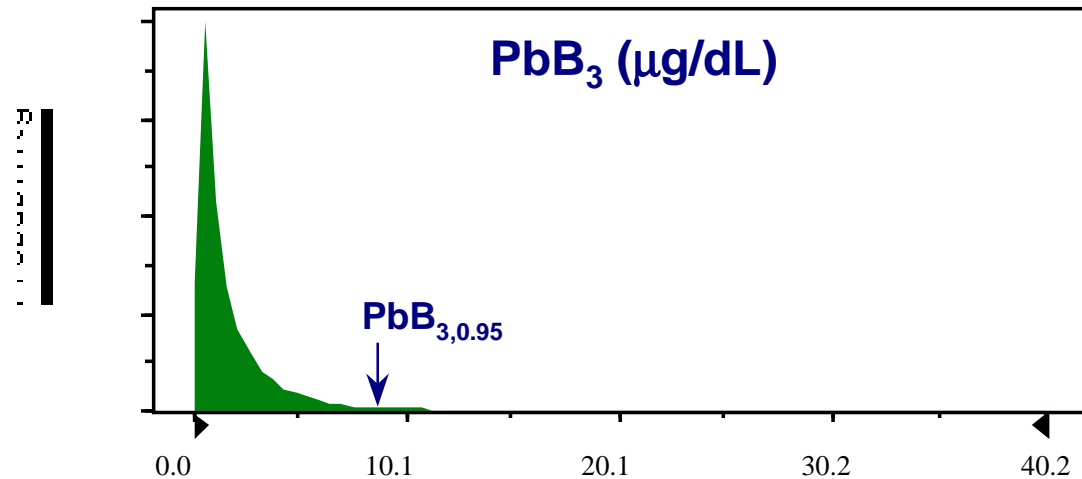
Measured by Contribution to Variance

Schematic of Exposure Estimates for Unbounded Simulated Population

What Is a Valid High-End Risk Estimate?

High-End Estimate

Bounding Estimate

Reasonable Worst-Case Estimate

$p_{0.50}$

$p_{0.90}$

$p_{0.95}$

$p_{0.98}$

$p_{0.99}$

$p_{0.999}$

Pb-Contaminated Smelter Site in Sandy, Utah

PbB$_3$ (μg/dL)

PbB$_{3,0.95}$

0.0	10.1	20.1	30.2	40.2

average	2 μg/dL
median	1.2 μg/dL
95th%ile	9 μg/dL
EPA estimate	17 μg/dL, > 98th%tile
	(potential bounding est.)
overestimation	1.9$\times$
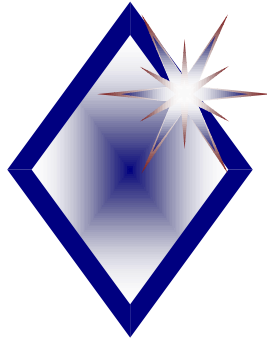
^{226}Ra-Contaminated Smelter Site in Soda Springs, Idaho
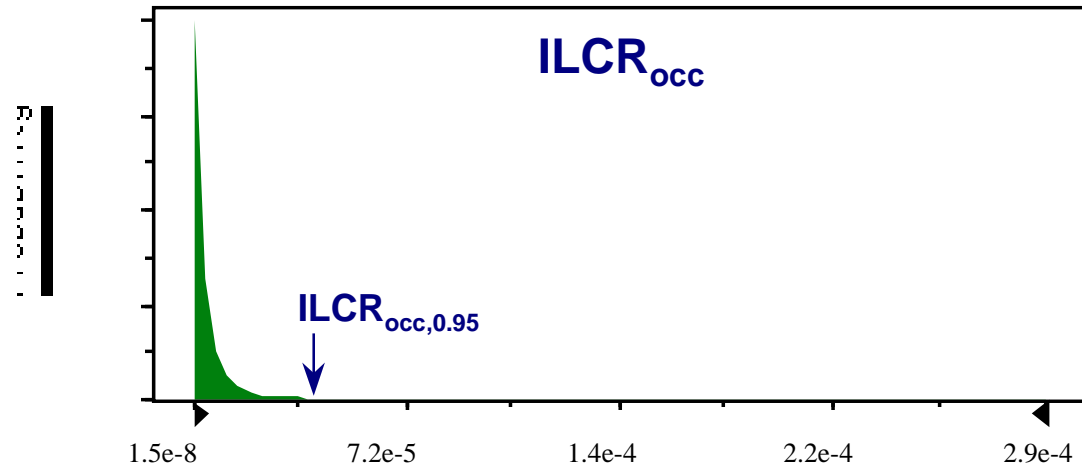
average 8×10^{-6} (8 in 1 million)

median 6×10^{-7} (6 in 10 million)

95th%ile 4×10^{-5} (4 in 100 thousand)

EPA estimate 2×10^{-3} (2 in 1 thousand),
               >> 99.9th%tile (bounding est.)

overestimation   50×
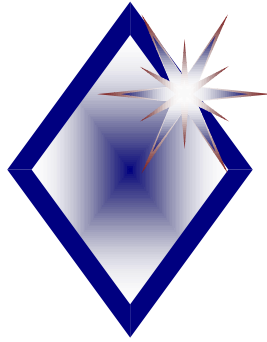
Deterministic vs. Probabilistic Modeling: Similarities

- ◆ **Similarities**
  - ◆ both approaches operate on the same fundamental model
  - ◆ both approaches often utilize the same data the difference being the amount of the data used (single point vs. distribution)
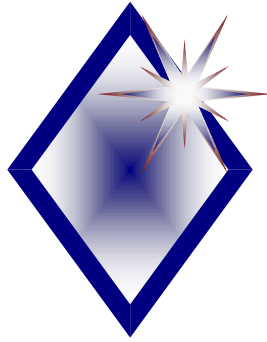
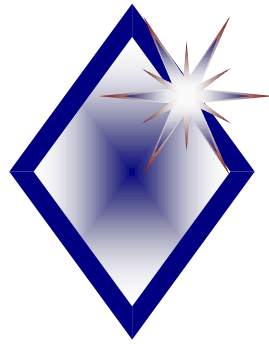# Deterministic vs. Probabilistic Modeling: Differences

- Differences
  - probabilistic approach utilizes complete distributions; deterministic approach utilizes single point from each distribution.
  - probabilistic approach quantifies uncertainty; deterministic approach does not.
  - probabilistic approach is generally can be more time and resource intensive than the deterministic approach.
 - probabilistic approach is capable of providing more realistic predictions; deterministic approach is more general.
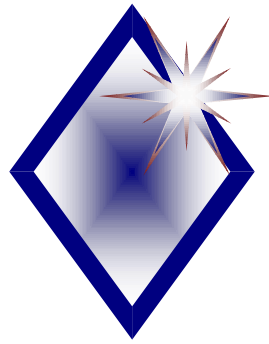
Probabilistic vs. Deterministic Modeling: Summary

◆ Deterministic modeling is relatively simple and less demanding of time and resources

◆ Deterministic modeling is a good screening tool

◆ Probabilistic modeling can be more realistic and helps to quantify uncertainty

◆ Monte Carlo simulation software and compatible hardware are readily available

Probabilistic vs. Deterministic Modeling: Summary

◆ To ensure the usefulness of the results of a probabilistic risk analysis, the following issues should be considered:

  ◆ Relevant and representative source of input data

  ◆ Selection of relevant subsets of data rather than aggregating all available data

  ◆ Consideration of quality of data at tails of input distributions

 ◆ Correlation of inputs and use of sensitivity analysis
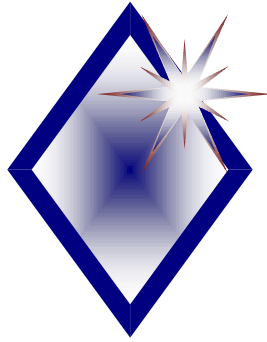
Probabilistic vs. Deterministic Modeling: Summary

◆ While probabilistic risk modeling is inherently more realistic, the quality of the outcome is directly related to the quality of the model structure and to the quality of the input assumptions ("garbage in garbage out," "garbage in gold out")
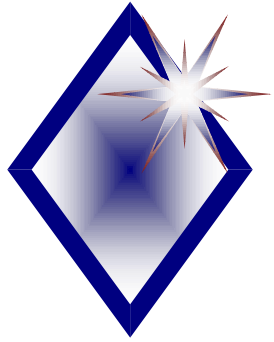
Objectives for this Session

◆ Overview

◆ Deterministic vs. Probabilistic Modeling

◆ Principles for use in MC Simulation

◆ On-going PRA activities

◆ Conclusion
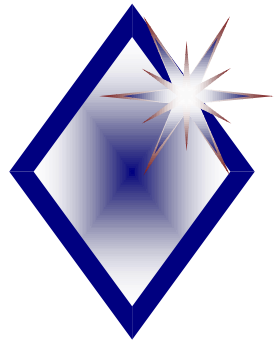
EPA Policy and Guidance

- EPA Policy for the Use of Probabilistic Analysis in Risk Assessment (May 1997)

- EPA Guiding Principles for Monte Carlo Analysis (March 1997)

- EPA Supplemental Guidance to RAGS: Guidance on Use of Probabilistic Risk Analysis in Risk Assessment, Vol 3, Part A (Draft, November 1999)
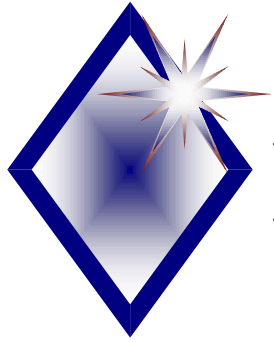
Existing EPA Policy

- Policy for Use of PRA in Risk Assessment
  - Use of MC or other such techniques in risk assessments shall not be cause, per se, for rejection of risk assessment by the Agency
  - PRA is not intended to apply to dose response evaluations for human health risk assessment
  - For ecological risk assessment the policy applies to all aspects including stressor and dose-response assessment
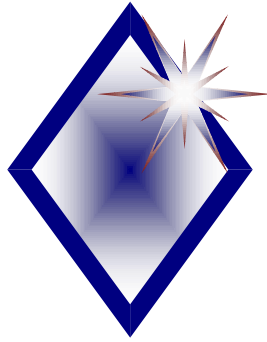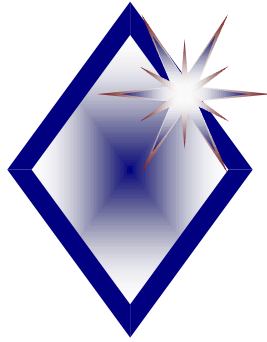
Eight Conditions for Use of Probabilistic Risk Assessment

1. Articulate purpose and scope. The questions the assessment tries to answer should be discussed and the assessment endpoints should be well defined.

2. Document methods used to allow reproduction of results. The methods used for analysis--all models, all data, all significant assumptions-- should be clearly documented and easily located in the write-up.
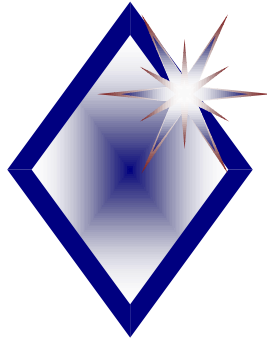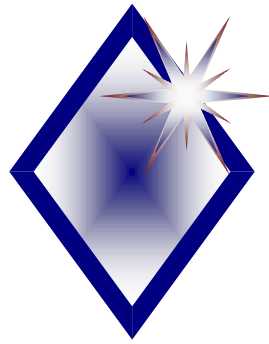
Eight Conditions for Use of Probabilistic Risk Assessment

3. <u>Present the sensitivity analyses</u>. The results of sensitivity analysis should be presented and discussed in the report, and as determined by sensitivity analysis, probabilistic techniques should be applied to the compound, pathways, and factors of importance to the assessment.

4. <u>Discuss correlations between input variables</u>. When strong correlations or dependencies between the input variables are either present or absent, they should be discussed along with the effects the correlations have on the risk assessment's output distribution.

Eight Conditions for Use of Probabilistic Risk Assessment

5. <u>Present and justify input and output distributions</u>. Detailed information for each input and output distribution should be provided in the report, including tabular and graphical representations of the distributions that indicate the locations of any point estimate of interest (mean, median, etc).

6. <u>Present numerical stability of output distributions</u>. The numerical stability of the central tendency and the tails of the output distributions should be investigated and discussed.
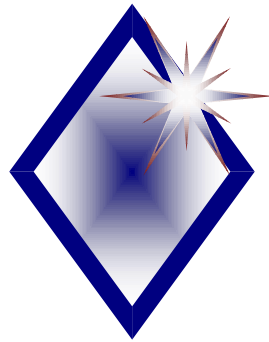
Eight Conditions for Use of Probabilistic Risk Assessment

7. <u>Provide point estimate results for comparisons</u>. Deterministic exposure calculations and risk estimates may be used to allow comparisons between the probabilistic analysis and past or screening level risk assessments and "is strongly encouraged." Scenario-specific questions can be answered using deterministic estimates, and such estimates can be used to facilitate risk communication.

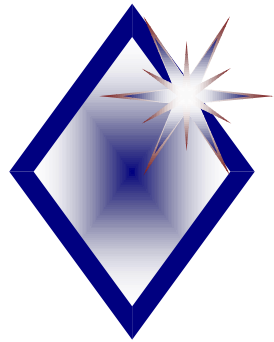# Eight Conditions for Use of Probabilistic Risk Assessment

8. Align exposure estimates with toxicity metric.

   Fixed exposure assumptions, such as body weight and exposure duration, are sometimes "embedded in toxicity metrics" (cancer potency factor, Reference Dose, etc.). The alignment of the probabilistic output distribution's exposure estimates with the toxicity metric should be discussed.
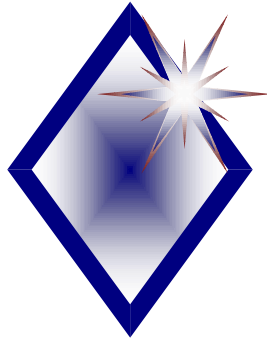
Possible Outline: RAGS Volume 3, Part A

- Introduction

- Definitions and Basics of the MC Method

- Advantages and Disadvantages of Deterministic and Probabilistic Risk Assessment

- When to Use Probabilistic Risk Assessment

- General Principals and Caveats of Probabilistic Analysis

- Preparation of the Work Plan and Review of a Probabilistic Risk Assessment

Additional Organizations with PRA Guidance

- Guidance for Use of Probabilistic Analysis in Human Health Risk Assessments (Jan 98)
  - Oregon Department of Environmental Quality
- Guidance for Submission of Probabilistic Exposure to the Office of Pesticide Programs' Health Effects Division (Feb 98)
 - Office of Pesticide Programs, USEPA
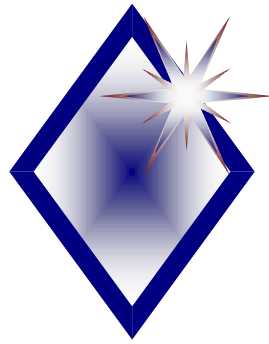
Objectives for this Session

- Overview
- Deterministic vs. Probabilistic Modeling
- Principles for use in MC Simulation
- On-going PRA activities
- Conclusion
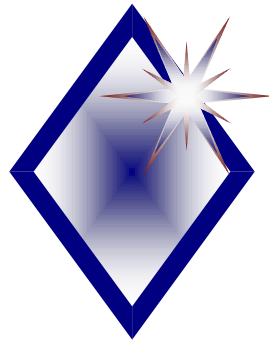
Risk Assessment Activities by the USAF

- Phased Approach used at Cape Canaveral
- Lawrence Livermore Phase I and II
- Probabilistic Risk Assessment Handbook
- Military Specific Exposure Factors
- Enhanced Site-Specific Risk Assessment
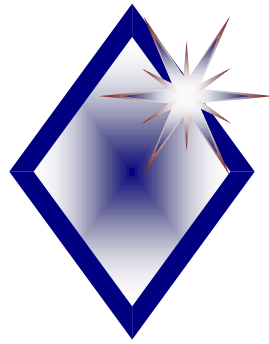
Phase Approach -- Cape Canaveral

◆ Goal: Focus resources on the most significant sites based on potential risk

◆ Overall risk analysis is divided into four levels of investigation

  ◆ Generic Risk-Based Screening (Phase 1)

  ◆ Site Specific Risk-Based Screening (Phase 2)

  ◆ Comprehensive Risk Assessment (Phase 3)

 ◆ Alternative Risk Assessment (Phase 4) - *Enhanced Site Specific Risk Assessment*
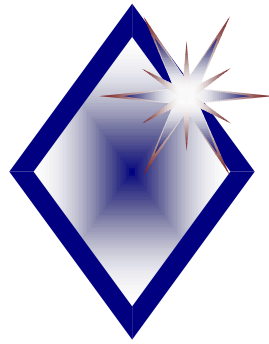
Phased Approach-- LLNL, CALEPA, DOD

- LLNL Phase I - Perform a stochastic (probabilistic) analysis of exposure and risk associated with TCE-contaminated groundwater at selected DOD sites in California--Beale AFB

- LLNL Phase II - Development of distributions of dose-response parameters related to the assessment of carcinogenic risk and noncarcinogenic hazard associated with TCE exposures.
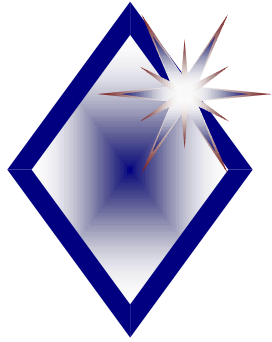
Additional Air Force Related PRA Activities

- ◆ AL Technical Report (February 1998)
  - ◆ Train AF risk assessors and risk managers
  - ◆ Provide introduction on techniques to evaluate uncertainty in risk assessment
- ◆ Military Specific Exposure Factors (1999)
  - ◆ Provide site-specific exposure distributions
 - ◆ Enhance PRA and reduce uncertainty
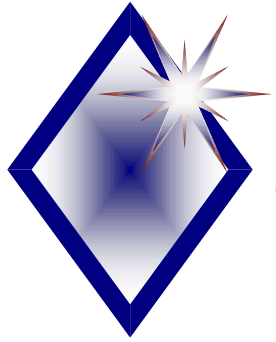
Objectives for this Session

- Overview
- Deterministic vs. Probabilistic Modeling
- Principles for use in MC Simulation
- On-going PRA activities
- Conclusion
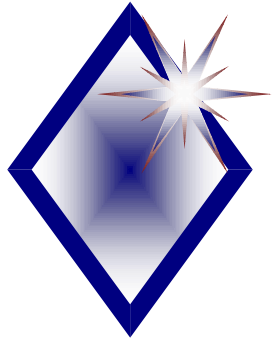
Common Mistakes to Avoid

- Use of point estimates for uncertain variables
- Use of distributions for non-variable parameters (e.g., averaging time)
- Underestimation of uncertainty
- Improper truncation
- Undocumented assumptions
- Confounding risk assessment/risk management
- Neglecting or misinterpreting background
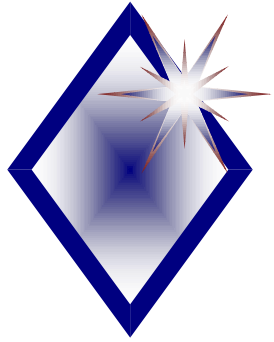
Common Mistakes to Avoid

- Awaiting perfection
- Failure to screen
- Failure to put estimated risks into perspective
- Failure to assess uncertainty in a meaningful way
- Failure to view risk assessment as an iterative process, and to use interim results as a project management tool
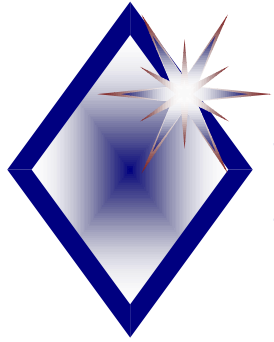
Barriers, Obstacles and Hurdles

◆ Managers need to see the value added

-what are the resource implications?

-will the use of PRA help or hinder our legal position? (EJ and current cleanup levels)

◆ Growing Pains

-absence of technical guidance & consensus on good ex's

-lack of knowledge by RPM and decision makers in its use

-how to communicate findings to the public

-surrogate data, defaults distributions, subjective uncertainty PDFs, distribution tails, rare events

-separation of true variability from uncertainty
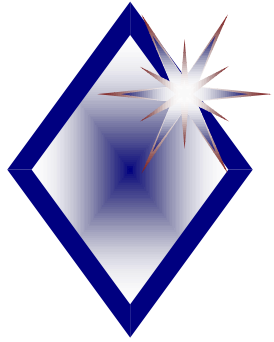
Summary

- Probabilistic risk assessment

  - Powerful tool for evaluating uncertainty; value is a function of knowing when, where, and how to use it

  - Risk managers are better able to leverage resources with an understanding of the uncertainties and distribution of possible risk associated with the site

  - Future guidance provided by EPA on its use at restoration sites

  - Need to apply science-based risk assessment to USAF restoration sites especially for LTM
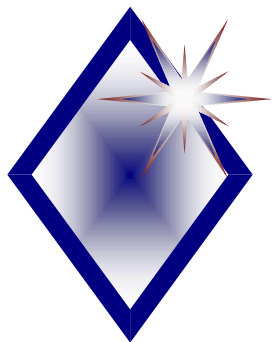
THANK YOU!

Captain Brian L. Sassaman

Comm (210) 925-3100 x 220

Email brian.sassaman@kelly.af.mil